

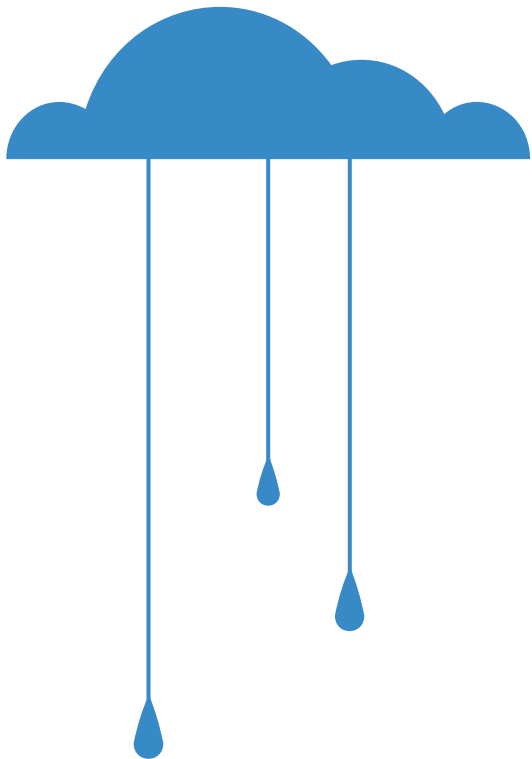


Rain in Australia Prediction Using Classification

By :

Afnan Alshehri
Kholod Alshehri





Introduction

Rainfall in Australia is extremely varied, owing to the region's large-scale atmospheric and oceanic causes.

In this project, we train classification models on the features to predict next-day rain .



Problem statement

Predict next-day rain by training classification models on the features .

Data Description This dataset contains about 10 years of daily weather observations from many locations across Australia.

Observations were drawn from numerous weather stations.

The daily observations are available from <http://www.bom.gov.au/climate/data>.



Dataset

Size

Before cleaning 145460 rows and 23 columns.

After cleaning 56420 rows and 22 columns.

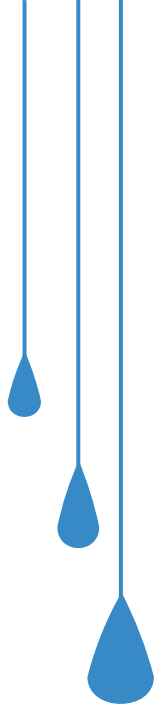
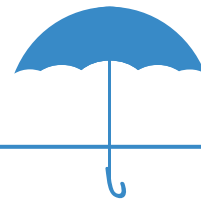
Source

From [kaggle](#)





Tools



Methodology



Problem understanding

01

Data loading

02

EDA

03

Results & insights

04

Baseline model

05

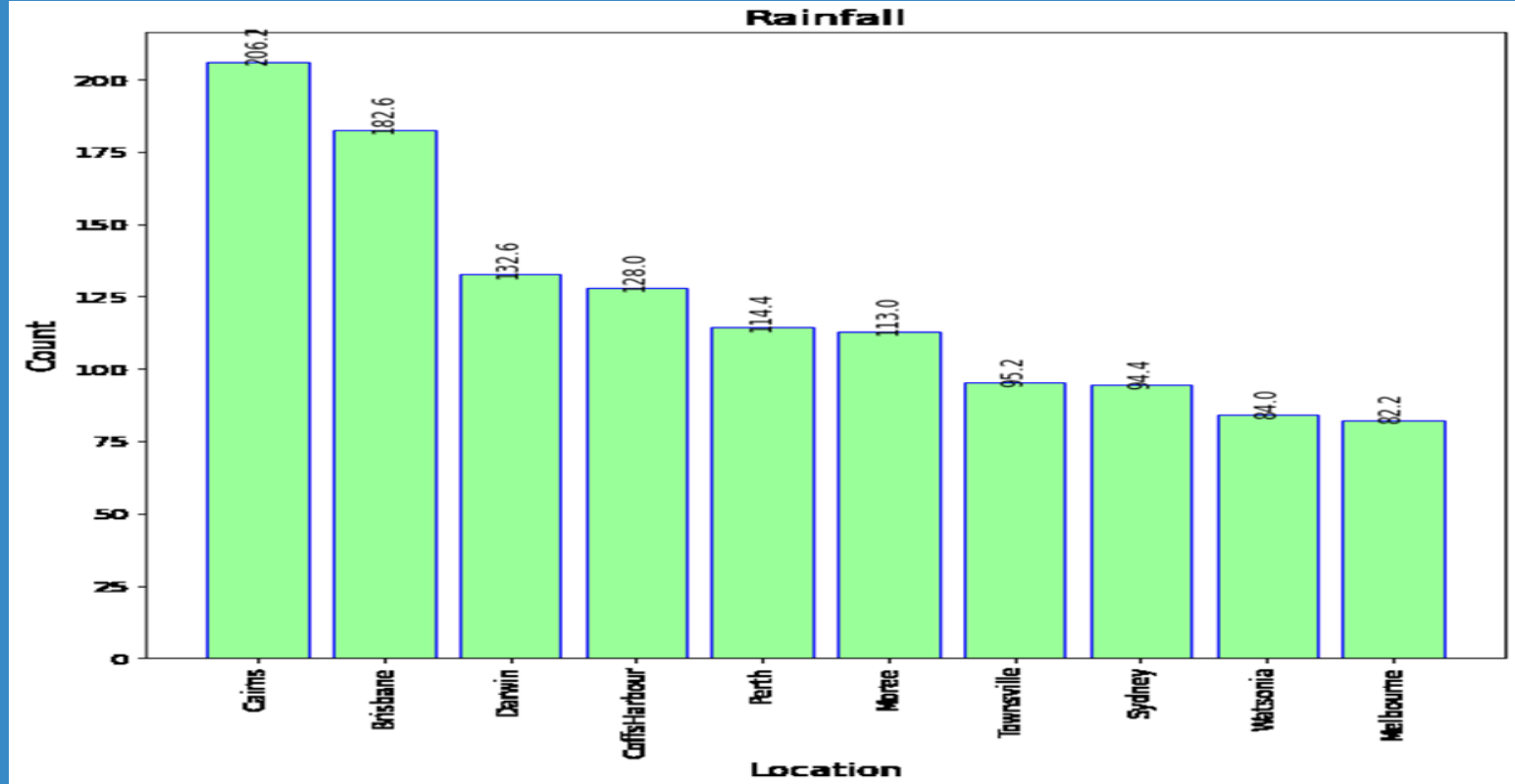
Experiments

06

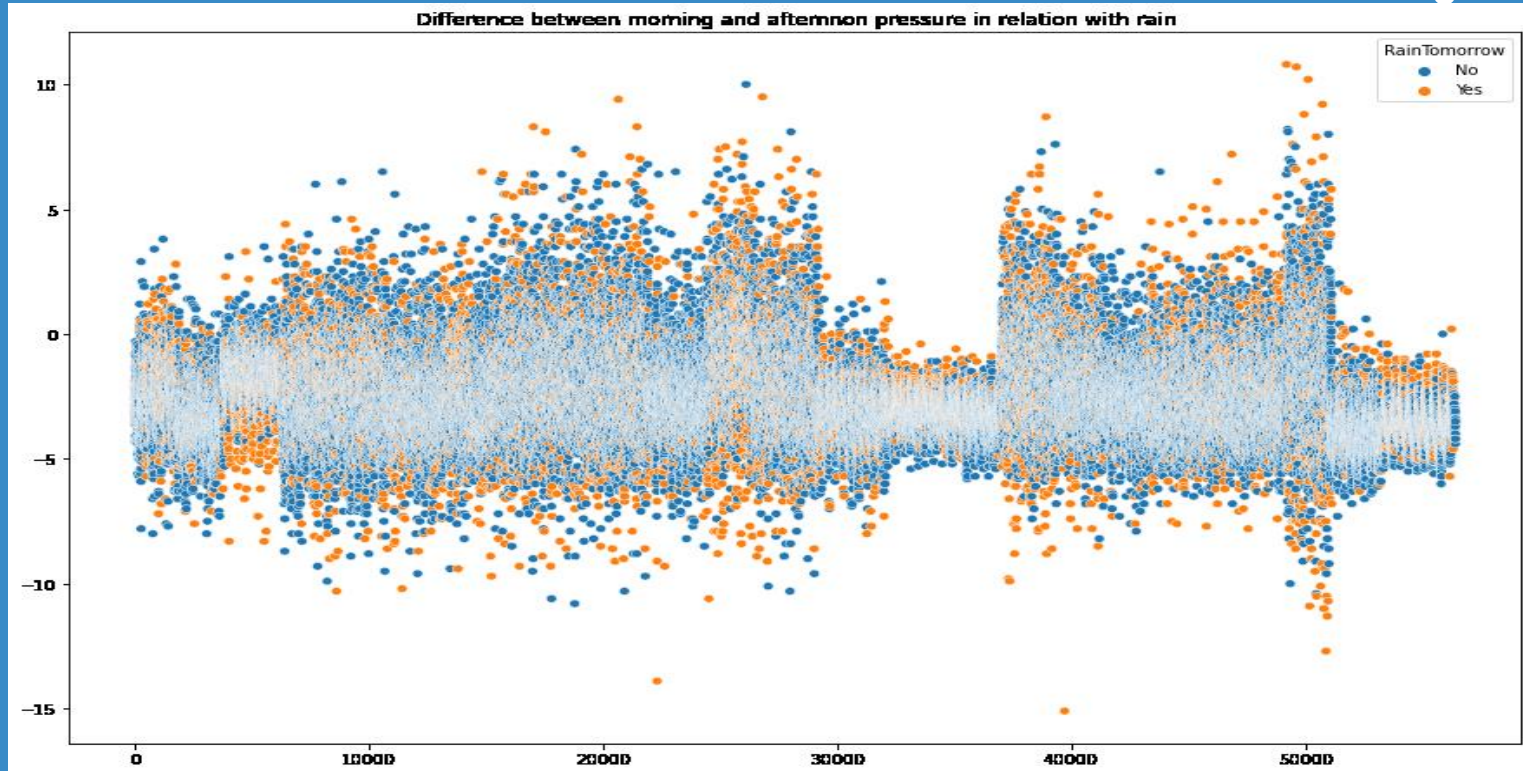
Classification

07

the country with most rain is : cairns



pressure not affecting the rain

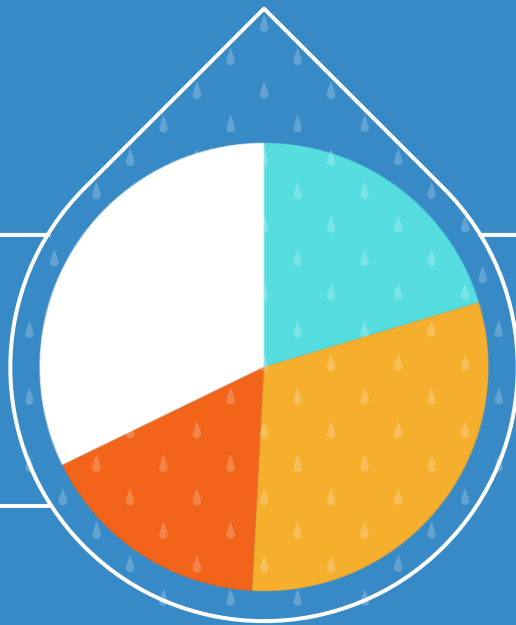


Data Splitting

90%
Train

5%
Test

5%
Validation



Baseline Model

Logistic
Regression

Before outliers

0.850

Train accuracy

0.853

Validation accuracy

After outliers

0.853

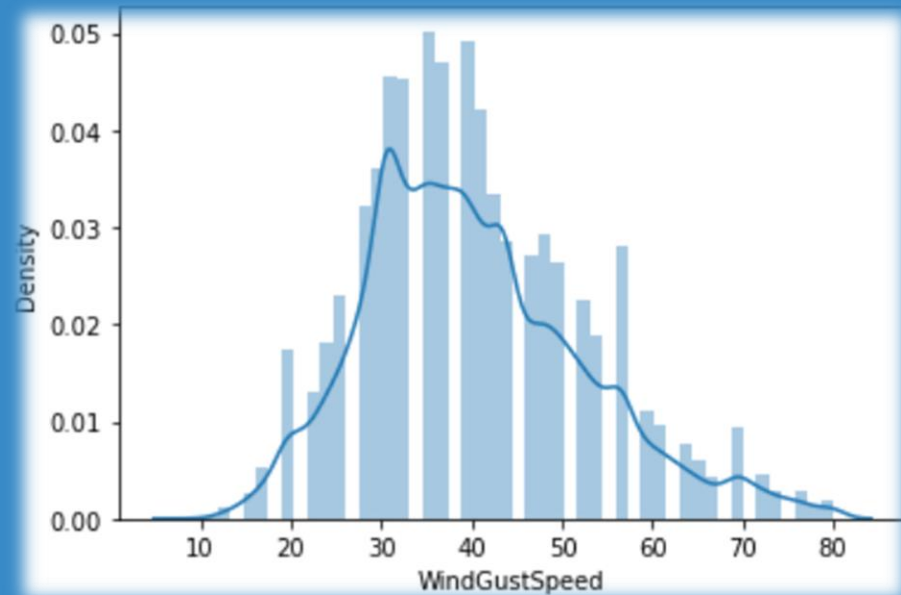
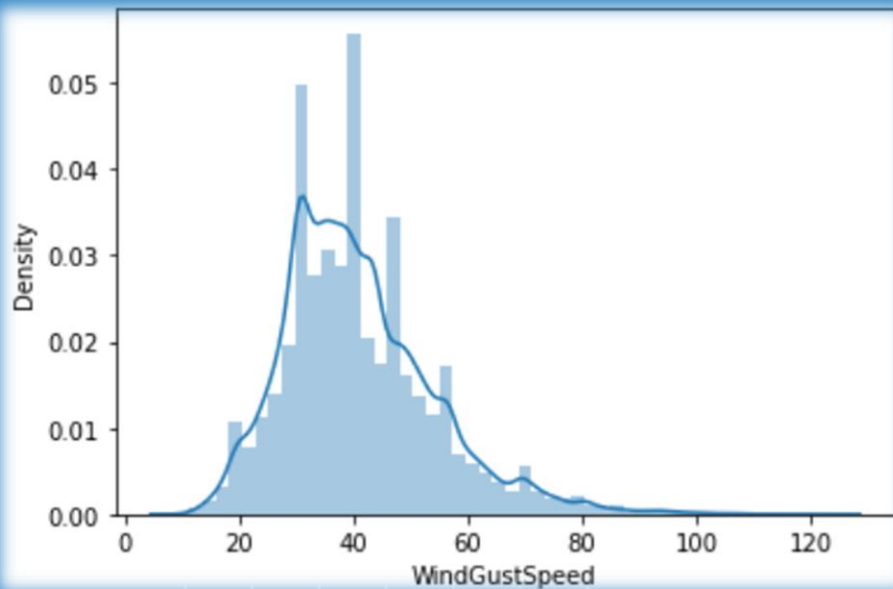
Train accuracy

0.860

Validation accuracy



Outliers



Model Scores

Classifier	Accuracy	Recall	Precision	F score
Logistic Regression	0.851	0.680	0.604	0.640
K-Nearest Neighbor	0.902	0.680	0.604	0.640
Naive Bayes	0.809	0.551	0.711	0.621
Decision Tree	0.861	0.720	0.536	0.637
SVM	0.840	-	-	-
XGboost	0.858	-	-	-
Random forst	0.853	0.604	0.604	0.625



Stacking Classifier

Classifier	Score
Stacking	0.867





Thank You
For Listening